



Mechanism of DNA primer synthesis by human PrimPol

Luis Blanco^{*,†}, Patricia A. Calvo, Alberto Diaz-Talavera,
Gustavo Carvalho, Nieves Calero, Ana Martínez-Carrón,
Cristina Velázquez-Ruiz, Sabina Villadangos, Susana Guerra,
María I. Martínez-Jiménez[†]

Centro de Biología Molecular Severo Ochoa, CSIC-UAM, Madrid, Spain

*Corresponding author: e-mail address: lblanco@cbm.csic.es

Contents

1. The priming problem and solutions	290
2. The AEP superfamily of primases	292
2.1 The human RNA primase is associated with Pol α	294
2.2 Human cells do have a PrimPol	294
2.3 Human PrimPol cellular context	295
2.4 Human PrimPol reprimers the leading strand to rescue stalled replication forks	297
3. Mechanism of DNA primer synthesis by human PrimPol	299
3.1 Eukaryotic PrimPols are monomeric, but with an extended AEP core	299
3.2 Sequential steps during primer synthesis by human PrimPol	299
4. Primase and polymerase configurations of human PrimPol	303
5. TLS abilities of human PrimPol are needed during DNA priming	305
6. AEP superfamily: Domain associations define functional classes	306
References	307

Abstract

PrimPol is the second primase discovered in eukaryotic cells, whose function is to restart the stalled replication forks during both mitochondrial and nuclear DNA replication. This chapter revises our current knowledge about the mechanism of synthesis of DNA primers by human PrimPol, and the importance of its distinctive Zn-finger domain (ZnFD). After PrimPol forms a binary complex with ssDNA, the formation of the pre-ternary complex strictly requires the presence of Mn²⁺ ions to stabilize the interaction of the incoming deoxynucleotide at the 3'-site. The capacity to bind both ssDNA template and 3'-deoxynucleotide was shown to reside in the AEP core of PrimPol, with ZnFD being dispensable at these two early steps of the primase reaction. Sugar

[†] Shared co-senior authorship.

selection favoring dNTPs versus NTPs at the 3' site is mediated by a specific tyrosine (Tyr¹⁰⁰) acting as a steric gate. Besides, a specific glutamate residue (Glu¹¹⁶) conforming a singular A motif (DxE) promotes the use of Mn²⁺ to stabilize the pre-ternary complex. Mirroring the function of the PriL subunit of dimeric AEP primases, the ZnFD of PrimPol is crucial to stabilize the initiating 5'-nucleotide, specifically interacting with the gamma-phosphate. Such an interaction is crucial to optimize dimer formation and the subsequent translocation events leading to the processive synthesis of a mature DNA primer. Finally, the capacity of PrimPol to tolerate lesions is discussed in the context of its DNA primase function, and its potential as a TLS primase.



1. The priming problem and solutions

The general process of DNA replication, which ensures the transmission of genetic information through generations, is conserved in all living organisms, but mechanistic details, molecular partners and interactions have diverged between prokaryotes and eukaryotes. DNA polymerases are central enzymes in this process, as they catalyze the faithful copy of each DNA strand during DNA replication. However, most DNA polymerases, irrespective of their specific role in replication, repair, damage tolerance, or variability, are unable to initiate DNA synthesis, as they cannot use a dNTP as the initial/priming source of the hydroxyl group required to transfer elongating nucleotides during DNA replication [1,2]. Nature has developed different solutions to solve this requirement of DNA polymerases, the need for a primer. As summarized in Fig. 1A, nicks, gaps, or DSBs, produced by damaging agents or dedicated nucleases, can act as eventual primers that will be used for specialized DNA polymerases involved in replication, repair and variability. The 3'-end of a stalled replicating strand can also serve as a primer for translesion synthesis (TLS) polymerases. On the other hand, some DNA polymerases can initiate DNA replication by using special, non-conventional primers as tRNA (reverse transcriptases), or terminal proteins (DNA polymerases from ϕ 29-like phages, adenoviruses, and some linear phages) which provide the 3'-OH of a specific amino acid residue [2,3]. Thus, in the latter examples, the very same enzyme performs the initiation reaction and the subsequent elongation of the DNA chain.

RNA polymerases appear to be exempt from such a priming problem, but strictly speaking, RNA polymerases do also require a non-conventional primer to initiate RNA synthesis (Fig. 1A and B): a single NTP, whose 3'-OH acts as the attacking nucleophile on a second NTP to form a

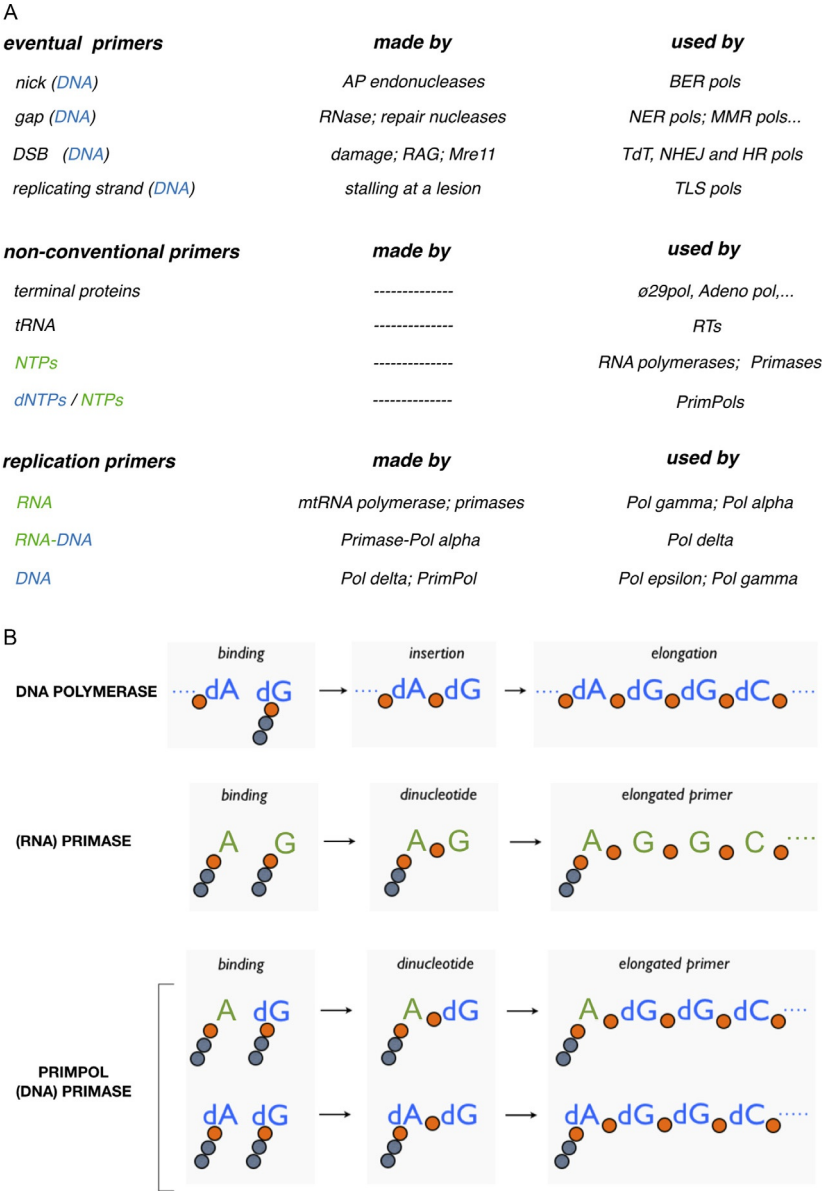


Fig. 1 The priming problem and solutions. (A) Different type of primers used to initiate DNA (and RNA) synthesis; (B) schematic representation of DNA polymerase-mediated DNA elongating versus primase and PrimPol-catalyzed initiation of RNA or DNA synthesis. Circles indicate the phosphates of a given deoxynucleotide/ribonucleotide, or the one involved in a phosphodiester bond. The α phosphate, the only remaining in the elongated product is colored in orange. The β and γ phosphates (in gray), are released as P_{PP}i during elongation, but remain at the beginning of the primers made by primases or PrimPols.

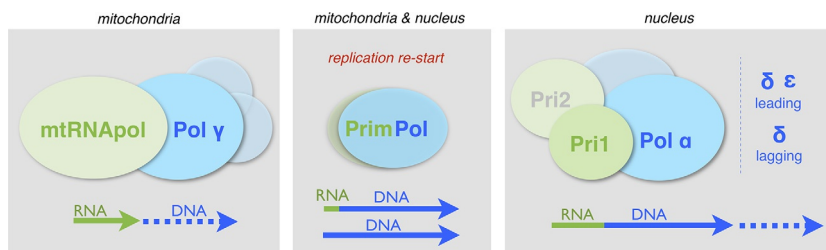


Fig. 2 Three priming alternatives operating during eukaryotic DNA replication. Left: mtRNA polymerase primes mtDNA replication by Poly. Middle: PrimPol generates DNA primers to re-start both mitochondrial and nuclear DNA replication under replicative stress. Right: The Pol α -primase tetramer makes hybrid RNA/DNA primers to fuel lagging and leading strand replication by Pol δ and Pol ϵ .

dinucleotide. Because of that, the mitochondrial RNA polymerase can synthesize abortive transcripts at the two mtDNA replication origins, which serve as DNA replication primers to be subsequently elongated by DNA polymerase gamma (Poly) [4] (see Fig. 2). In fact, the most common priming mechanism for DNA replication relies on a specialized RNA polymerase, called primase, which is able to initiate synthesis from a single ribonucleotide triphosphate (NTP) to produce short RNA primers whose 3'-OH can be elongated by replicative DNA polymerases, as DNA polymerase alpha (Pol α) in eukaryotic cells [1,5,6]. Interestingly, a preference for a DNA primer distinguishes the other eukaryotic DNA polymerases involved in DNA replication. Thus, the main eukaryotic replicase DNA polymerase delta (Pol δ) can use the hybrid RNA/DNA primers made by Pol α -primase complex to copy the lagging strand, and also initiate synthesis of the leading strand [7]. DNA polymerase epsilon (Pol ϵ) will continue synthesis of the leading strand by processive extension of the DNA primer provided by Pol δ . Strikingly, as it will be further discussed in this chapter, PrimPol is a new type of primase (DNA primase), which can use dNTPs as non-conventional primers, to generate a valid DNA primer for Poly and Pol ϵ (Fig. 1B).



2. The AEP superfamily of primases

Primases are classified into two main families that differ in their structure as well as in their interactions with other proteins of the replication complex. Based on their primary sequence structure, primases can be divided into two groups: the DnaG superfamily with a TOPRIM (topoisomerase-primase) fold domain [8], which contains bacterial and

bacteriophage primases; the AEP (archeo-eukaryotic primase) superfamily, which contains an RRM (RNA recognition motif) fold domain [9,10]. AEPs include heterodimeric archaeal and eukaryotic primases, but also monomeric enzymes named PrimPols, that can have a helicase fused to the AEP core. As PrimPols have been characterized in Archaea and Eukarya, but also in Bacteria, it has been suggested that AEPs were recruited at the base of the evolution of the archaeo-eukaryotic lineage, with a subsequent acquisition by bacteria via horizontal gene transfer [11]. Based on amino acid sequence similarity, AEPs were classified into three higher-order clades:

1. *The AEP proper clade*, which includes the small subunits of archaeal and eukaryotic primases (classical AEPs), the Lef-1 baculoviruses primases, and the bacterial NHEJ polymerases. Classical AEPs and Lef-1 family form a functional complex with their respective large, non-catalytic subunit [12]. On the other hand, bacterial AEPs specialized in NHEJ appear to be fused to a DNA ligase and a nuclease. This last family is a special case of primase-like enzymes lacking the ability to initiate *de novo* synthesis to be specialized in processing the ends of double-strand DNA breaks, thus contributing to the prokaryotic NHEJ system [13].
2. *PrimPol-like clade*. This clade includes seven distinct families of primases from bacteria, archaea, viruses, bacteriophages, and plasmids. Primases in this clade often have one of the following two domains: Primase-C Terminal-1 and Primase-C Terminal-2 (PriCT-1, PriCT-2). The first described PrimPol was Orf904 from *Sulfolobus islandicus* archaeal plasmid pRN1, which is a highly compact multifunctional enzyme with ATPase, primase and DNA polymerase activity in its N-terminal region and a helicase activity in its C-terminal region [14]. This protein synthesizes a primer of approximately seven deoxyribonucleotides and possesses high sequence specificity for the initiation site [15]. Some recently characterized members of the AEP superfamily, as BcMCM from *Bacillus cereus* [16] and TthPrimPol from *Thermus thermophilus* [17], behave as PrimPols.
3. *The NCLDV-herpesvirus primase clade*. This clade includes predicted primases present in large nucleo-cytoplasmic DNA viruses (NCLDV), kinetoplastids, and a new subfamily of eukaryotic primases originally termed EukPrim2 [11]. A signature of this clade is the presence of a highly conserved C-terminal domain containing a Zn-finger motif. Based on the structural domain organization, this clade can be subdivided into two families: the iridovirus primase family, and the herpes-poxvirus primase family, the latter including the Eukprim2 subfamily, represented

by a putative second human primase [11]. As it will be described later, the human Eukprim2 was renamed PrimPol, as it has both DNA primase and DNA polymerase activities like pRN1 PrimPol.

2.1 The human RNA primase is associated with Pol α

Unlike monomeric bacterial primases, eukaryotic primases belonging to the AEP proper clade are heterodimers of catalytic (p49) and regulatory (p58) subunits. The small primase subunit (p49), which contains the active site for RNA primer synthesis, is often referred to as Pri1 or PriS. The large primase subunit (p58), which coordinates RNA primase and polymerase activities, is named Pri2 or PriL [5,11,18]. These primases form a heterotetrameric complex with the two subunits of DNA polymerase alpha (Pol α), having an apparent mass of more than 300kDa. This tetrameric complex, named the Pol α -primase complex, synthesizes the RNA primer, which initiates synthesis of the leading strand at each replication origin, and the multiple RNA primers required to synthesize the Okazaki fragments at the lagging strand during nuclear DNA replication [19]. The large essential subunit (p58/Pri2/PriL) participates in primer synthesis, counts the number of nucleotides in an RNA primer, assists the release of the RNA primer-template from p49/Pri1/PriS, and transfers it to the Pol α active site [20,18]. After a slow dinucleotide formation step, the Pol α -primase complex rapidly synthesizes RNA primers 7–10 nucleotides long. Extension of these RNA oligonucleotides into RNA–DNA primers of ~ 30 nucleotides by Pol α is required before their transfer to the active site of the more processive Pol δ and Pol ϵ (Fig. 2).

2.2 Human cells do have a PrimPol

In 2008 we learned about the existence of a putative second primase in human cells, initially named CCDC111 or EukPrim2, as Koonin's group had reported in a detailed *in silico* study of AEPs, the Archeo-Eukaryotic superfamily of Primases [11]. We promptly decided to start the characterization of this new enzyme, by expressing it in *Escherichia coli*. After several attempts, we succeeded in getting the protein soluble when fused to a 10xHis-tag at its amino-terminal end. After purification of the enzyme, its characterization as a conventional primase activity (synthesis of RNA primers) was unsatisfactory, as the purified fraction appeared to be endowed with some DNA polymerase activity, specifically activated by using Mn²⁺ ions. Construction of a catalytically-dead mutant, by eliminating two of

the putative metal ligands, demonstrated that such a DNA polymerase activity was not a contaminant, but intrinsic to CCDC111. These unexpected results guided us to explore the possibility that this putative new primase could be able to make *de novo* DNA primers. As expected, the primase was much more efficient in making DNA primers than RNA primers, also when activated by Mn^{2+} ions [21]. Interestingly, such a kind of dual enzyme endowed with both DNA primase and DNA polymerase activity had been previously reported in an archeal plasmid (pRN1 from *S. islandicus*) [22], and named Prim-Pol. Accordingly, we proposed the name PRIMPOL for the human gene and PrimPol for the human enzyme to the HUGO committee in October 2011, and our proposal was accepted prior to the publication based on the enzymatic properties that we showed *in vitro*. As conventional primases serve to “prime DNA replication”, they are frequently referred to as DNA primases; however, this nomenclature is misleading given the nature of the primers that they make (RNA) (see Fig. 1A and B). The discovery of PrimPol, a unique primase having deoxyribose sugar selectivity, further strengthens the need for a more precise terminology. Accordingly, conventional primases should be defined as RNA primases, whereas PrimPols are DNA primases *sensu stricto*.

A more extensive analysis of the primary sequence of human PrimPol confirmed that it belongs to the AEP superfamily, and identified 14 highly conserved motifs conserved among the members of this new PrimPol family [21]. Eukaryotic PrimPols are widely distributed throughout the kingdoms of fungi, animals and plants (Fig. 3). The gene of this new type of primase (PRIMPOL) appears to be acquired early in the eukaryotic evolution and subsequently lost in several occasions in fungi and animal kingdoms, such *Saccharomyces cerevisiae*, *Caenorhabditis elegans* and *Drosophila melanogaster*.

2.3 Human PrimPol cellular context

Cellular fractionation experiments have shown that PrimPol localizes to both the nucleus and mitochondria of human cells, being thus the first primase and second polymerase known to operate in mitochondria [21]. PrimPol provides a new and more ergonomic solution to the priming problem: the direct synthesis of DNA primers, unlike conventional primases that generate RNA primers, is a simplified alternative to the action of the Pol α -primase complex (Fig. 2). PrimPol primase activity was demonstrated to be essential to reinitiate stalled forks challenged by UV or hydroxyurea treatment, most likely in the leading strand, and even to maintain the normal replication rate in the

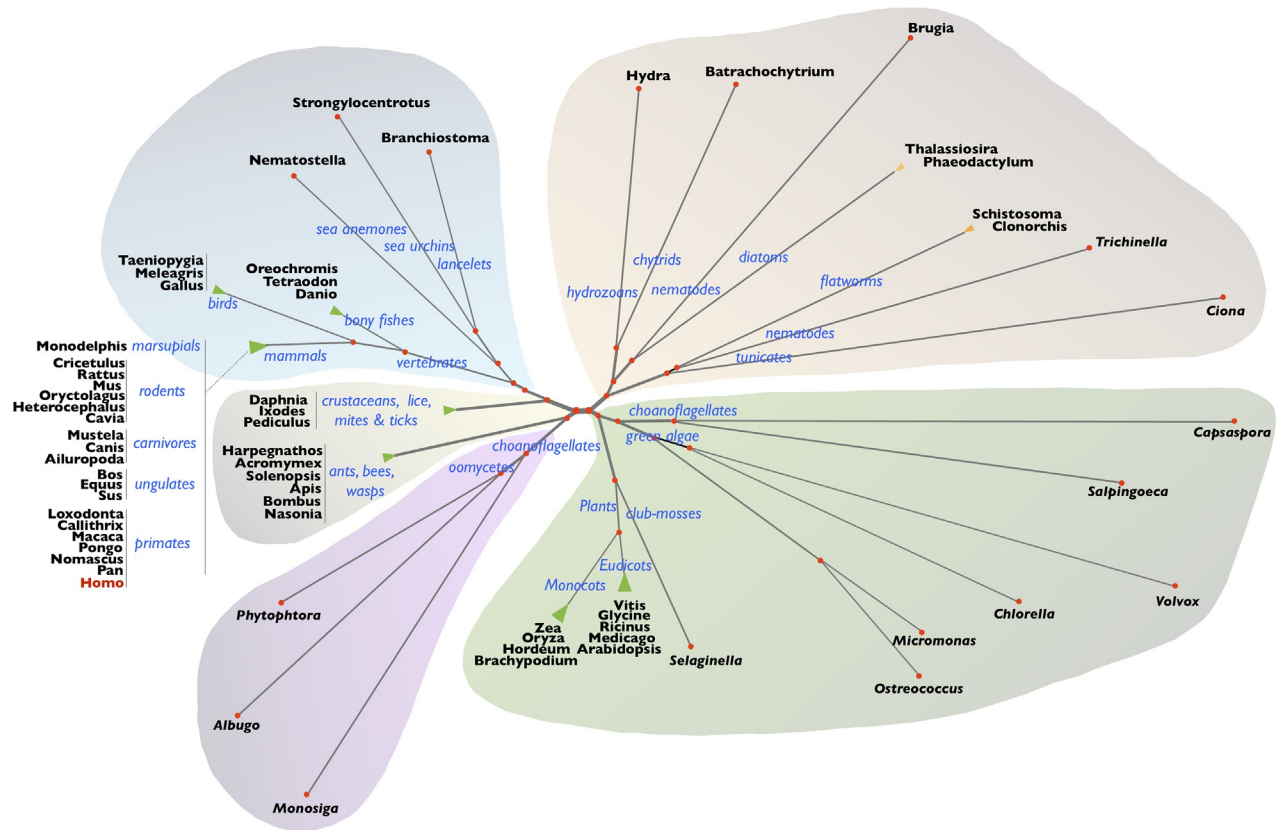


Fig. 3 Phylogenetic tree supporting the monophyletic origin or eukaryotic PrimPols.

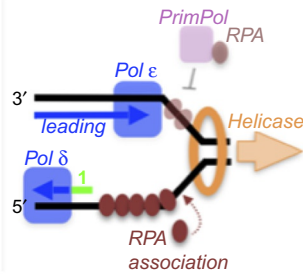
absence of exogenous stress during nuclear replication [23,24]. Besides, PrimPol primase activity was shown to be required also during mitochondrial DNA replication, by reinitiating synthesis after UV damage or in the presence of chain-terminating nucleotides [21,25].

2.4 Human PrimPol reprimers the leading strand to rescue stalled replication forks

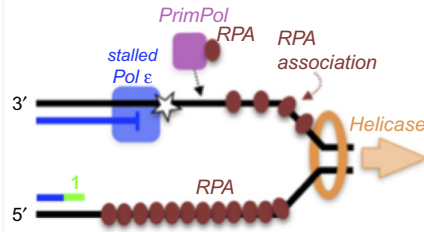
The enzymatic potential of human PrimPol as a DNA primase-DNA polymerase provides new TLS solutions to assist fork progression depending on the lesion or abnormal structure causing the stalling of the replication machinery: (1) PrimPol can synthesize DNA primers ahead of UV lesions, blocked primer-terminus, G-quadruplexes, or even R-loops to allow DNA replication to continue [23,26,27,28]; (2) PrimPol can realign primers ahead of “unreadable lesions” such as abasic sites and pyrimidine(6-4) pyrimidone photoproducts (6-4PPs), thereby skipping the lesion [29]; (3) PrimPol can incorporate nucleotides opposite DNA lesions as 8-oxo-7, 8-dihydro-2'-deoxyguanosine (8oxodG), like a regular TLS DNA polymerase [21,29,30].

During normal replication fork progression, the leading replicase and the helicase are coordinated to form the newly synthesized DNA strands, either continuously (leading strand) or in multiple steps that require recurrent priming events at the lagging strand (Fig. 4). The human single-stranded DNA binding protein, RPA, mainly binds to the uncopied lagging strand template; on the other hand, the length of ssDNA template in the leading strand is maintained relatively short preventing PrimPol repriming on this ssDNA (Fig. 4). Under replicative stress, the leading strand replicase Pol ϵ can be eventually stalled becoming uncoupled to the unwinding helicase, promoting the accumulation of a long stretch of uncopied ssDNA template, which could not be readily covered by RPA. Thus, PrimPol finds the opportunity to bind in proximity to the stalling site, likely via specific interactions with RPA [24,31,32] to reprime the leading strand and to trigger a polymerase switch with Pol ϵ for DNA primer elongation. In this way, PrimPol re-establishes the normal progression of the replication fork, including the synthesis of new Okazaki fragments generated at the lagging strand. The gap left behind in the leading strand can be mitigated by DNA damage tolerance events, such as direct lesion bypass by translesion synthesis machineries, or template switching mediated by the newly synthesized Okazaki fragment (Fig. 4).

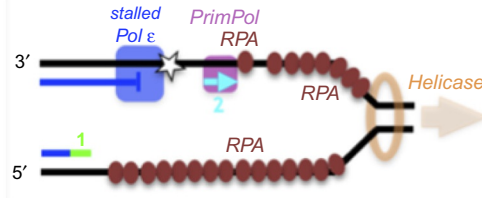
Normal replication



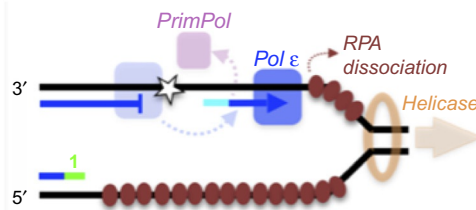
Stalling & replication stress



Repriming



Polymerase switch



Re-established fork progression

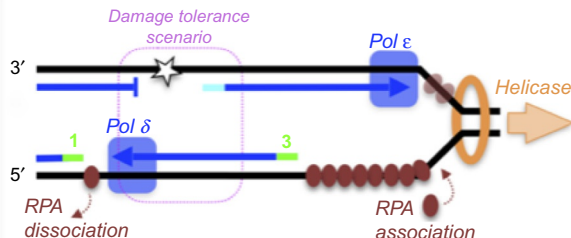


Fig. 4 Schematic representation of PrimPol acting at the stalled replication fork. After PrimPol recruitment via RPA and subsequent synthesis of a DNA primer, fork progression is re-established, and a lesion-containing gap is left behind, which requires the direct action of TLS polymerases, or template-switch (damage avoidance) events to copy the newly synthesized/undamaged daughter strand. RNA primers in the lagging strand (green) or an eventual DNA primer made by PrimPol in the leading strand (cyan) are numbered according to their proposed order of synthesis. Adapted from M.I. Martinez-Jimenez, A. Lahera, L. Blanco, Human PrimPol activity is enhanced by RPA, *Sci. Rep.* 7 (2017) 783.

Importantly, the discovery of PrimPol-mediated repriming reformulates the context of the posterior DNA tolerance/repair reactions, requiring future investigation on how the lesion-containing gaps left behind after PrimPol action are filled, processed and ligated. Recent findings demonstrated that most of the TLS events do occur after leading-strand repriming (reviewed in Ref. [33]).



3. Mechanism of DNA primer synthesis by human PrimPol

3.1 Eukaryotic PrimPols are monomeric, but with an extended AEP core

Conventional AEP primases are dimeric, and the small catalytic subunit (i.e., *HsPrim1/PriS*) has the three conserved motifs (A, B and C) that form the RNA primase active site (see Fig. 5A). By contrast, eukaryotic PrimPols are monomeric with an AEP core containing the same three conserved motifs, together with a C-terminal extension which contains a Zn-finger [21,24] and two twin-motifs involved in RPA interaction [24,31,34]. As already mentioned, PrimPol is recruited to the stalled fork by interaction with the RPA filament formed at the leading-strand ssDNA template (Fig. 4).

The Zinc-finger domain (ZnFD) of human PrimPol has three conserved cysteines and a histidine (Cys-His-Cys-Cys) with the potential to coordinate a Zn^{2+} atom, very similar to those in AEP primases from the herpesvirus family [11]. The ZnFD in human PrimPol was shown to be essential for the primase activity both *in vivo* and *in vitro*, but fully dispensable for the polymerase and TLS activities [23,34]. Expression of human PrimPol versions with a deletion of the ZnFD, or solely with single mutations at Zn^{2+} ligands, did not rescue normal fork rate and fork stalling in PrimPol-depleted cells, demonstrating the importance of the Zn^{2+} finger, and implying that the primase activity of PrimPol is required during nuclear DNA replication [23], and also during mitochondrial DNA replication [25].

3.2 Sequential steps during primer synthesis by human PrimPol

By using templates containing a preferred and single priming site, we were able to delineate the individual steps and substrate preference of PrimPol when building a primer, and to determine the relevance of the ZnFD at each individual step [35].

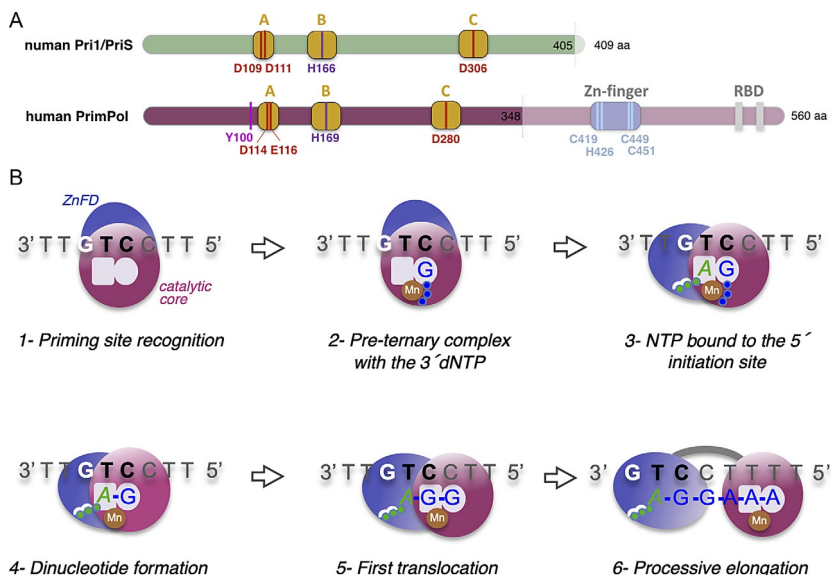


Fig. 5 Mechanism of DNA priming by human PrimPol. (A) Structural comparison of human Pri1/PriS/p49 and human PrimPol. In the linear representation shown above, the AEP cores (shown in dark color) contain the highly conserved motifs A, B and C with the invariant metal and nucleotide binding ligands. The sugar selector (Tyr¹⁰⁰) in human PrimPol is also indicated. The C-terminal extension of human PrimPol (shown in light pink) contains the Zn-finger and the bipartite RPA binding domain (RBD). (B) Sequential steps during primer synthesis by human PrimPol. (1) Priming site recognition: binding of PrimPol to ssDNA does not require the Zn-finger domain (ZnFD; dark purple). (2) Pre-ternary complex formation: the 3'-deoxynucleotide (dGTP; in blue) binds to the elongation site without the influence of the ZnFD, and requires manganese ions (a single Mn^{2+} sphere (ochre) is depicted for simplicity) and template base-complementarity. (3, 4) Nucleotide binding at the 5'-initiation site, and dinucleotide formation: the 5'-ribonucleotide (ATP; in green) binds to the initiation site to promote dinucleotide 3pA-G formation. The Zn-finger domain plays a key role in binding the triphosphate moiety of the 5'-site nucleotide, keeping a closed conformation that promotes catalysis. (5) First translocation: ZnFD maintains the interaction with the 5'-end triphosphate, facilitating dimer translocation and next nucleotide addition. (6) Processive elongation: the primer is elongated processively, likely favored by a persisting interaction of the ZnFD with the 5'-triphosphate at the end of the primer, and by its flexible tethering to the catalytic core of PrimPol.

3.2.1 Binary and pre-ternary complexes

Human PrimPol binds first the ssDNA template, preferentially at pyrimidine-rich motifs, and divalent metals are not required for the stabilization of this *enzyme:ssDNA binary complex* (Fig. 5B, step 1). After PrimPol binding to ssDNA and recognition of a preferred template sequence (GTC in our assays), PrimPol binds a first nucleotide (dGTP) at the elongation site,

which is selected by Watson–Crick hydrogen bonding with the templating C at the recognition sequence 3′GTC5′, thus forming a so-called *pre-ternary complex* (Fig. 5B, step 2). The ZnFD was shown to be dispensable at these two early steps of primer synthesis, the binding of both ssDNA template and 3′-nucleotide.

Importantly, the preternary complex is formed with the complementary 3′ site deoxynucleotide (dGTP), but not with GTP. Thus, the sugar discrimination observed at this step of the primase reaction explains the specificity of human PrimPol for making DNA primers instead of RNA primers. Recent kinetic studies demonstrated the existence of a steric gate to discriminate against ribonucleotides, and that a single residue (Tyr¹⁰⁰; see Fig. 5A) acts as a sugar discriminator in human PrimPol [36]. Interestingly, a mutation of this residue to Histidine (as found in the catalogue of somatic mutations in cancer, COSMIC) disables the steric gate of human PrimPol, enabling it to start and elongate primers efficiently with NTPs. This was shown to be due to the improved binding of NTPs at the 3′ site, evidenced during the formation of the pre-ternary complex. Given the higher concentration of NTPs in the cell, the cancer-associated variant Y100H becomes a more efficient primase which confers both mouse and human cells with increased resistance to hydroxyurea, suggesting a potential advantage to minimize the replicative stress associated to dNTP deprivation during cancer growth [36].

The pre-ternary complex formation strictly requires the presence of Mn²⁺ ions to stabilize the interaction with the incoming nucleotide at the 3′-site. This requisite is also in perfect agreement with the Mn²⁺-dependent activation of the DNA primase activity of human PrimPol [21]. A crucial aspect of the “two-metal-ion catalytic mechanism” of polymerization shared by polymerases and primases is the precise interaction of two (identical or different?) metal ions with the three conserved catalytic carboxylates to allow the proper alignment of the attacking nucleophile and the alpha-phosphorus atom of the incoming nucleotide. We have recently shown that substitution of each metal ligand of human PrimPol (Asp¹¹⁴, Glu¹¹⁶, Asp²⁸⁰) with alanine does not affect binding to ssDNA, but abolishes the stabilization of the 3′-nucleotide to form the PrimPol:ssDNA:dNTP pre-ternary complex [37]. Importantly, one of these metal-ligand carboxylates is invariantly a glutamate (Glu¹¹⁶) residue in eukaryotic PrimPols (motif DxE; see Fig. 5A), unlike the more conventional aspartate found in other primases and polymerases. A mutation of Glu¹¹⁶ to Asp in human PrimPol strongly affected the formation of the pre-ternary complex, especially in the presence of both Mg²⁺ and Mn²⁺. Recently, the crystal structure of the

catalytic core of human PrimPol has been solved in complex with a template/primer, one metal ion (Ca^{2+}) which occupies the metal B position, and dATP as the incoming 3'-site nucleotide, albeit lacking the second metal ion (A) at the active site [38]. Thus, in the absence of more precise structural information, we speculate that the distinctive Glu¹¹⁶ residue (motif DxEx) of human PrimPol may coordinate Mn^{2+} as metal ion B, and interacts with the α and γ phosphate groups of the 3'-site incoming nucleotide, thus promoting the use of Mn^{2+} to stabilize the pre-ternary complex [37].

3.2.2 Binding the 5' nucleotide and dimer formation

The next steps after pre-ternary complex formation are the binding and selection of the initiating nucleotide, which will occupy the PrimPol primer/initiation 5'-site, followed by catalysis and subsequent dimer formation. Attempts to demonstrate the formation of an alternative pre-ternary complex with only the initiating nucleotide (ATP) or even a quaternary complex (enzyme:ssDNA:ATP:dGTP; Fig. 5B, step 3) failed, supporting previous evidence that primases weakly/transiently interact with the initiating nucleotide that will become the 5'-end of the growing primer. Therefore, 5'-nucleotide binding had to be inferred from the analysis of the first catalytic event, i.e., dimer formation (Fig. 5B, step 4), as a function of the number of phosphates of the 5'-nucleotide provided. Human PrimPol largely preferred a triphosphate in the initiating nucleotide, followed by a diphosphate, and monophosphate as a much less efficient nucleotide substrate [35]. A similar preference for triphosphate-containing initiating nucleotides has been recently described for a PrimPol-like phage enzyme [39]. Strikingly, the preference for a triphosphate in human PrimPol is mediated by the ZnFD, explaining the reported reduction in primase activity of a ΔZnFD mutant [23]; however, the ZnFD was much less important for 5'-nucleotide substrates with one or two phosphates; indeed, in the absence of the ZnFD, ADP and dADP were even more efficient than ATP and dATP, respectively, supporting an inhibitory effect of the γ -phosphate. These data imply that the ZnFD of PrimPol is involved in the binding and selection of the 5'-site nucleotide, most likely by establishing interactions with the γ -phosphate moiety of triphosphate-containing nucleotides, making them both valid and optimal substrates for catalysis of the initial dimer. Our analysis also showed a preference for a ribose sugar in the 5'-site nucleotide. That preference was maintained when comparing nucleotides with two or three phosphates, and even in the absence of the ZnFD, suggesting that some residue(s) of the APE core of PrimPol can establish stabilizing interactions with the 2'-OH group of the initiating nucleotide [35].

3.2.3 Translocation and processive synthesis

After dimer formation, translocation is needed to copy the next nucleotide of the template (Fig. 5B, step 5). Upon translocation, the second nucleotide of the dimer has to move back to be repositioned at the primer/initiation site of the enzyme. Given the preference of PrimPol for inserting dNTPs at the elongation site, even during dimer synthesis [21], the configuration of the initiation site after the first translocation, and also during the next elongation events (Fig. 5, step 6), must be compatible with a deoxyribose, and not only with a ribose as in conventional RNA primases. Interestingly, the efficiency of translocation and further elongation steps also took advantage of a 5'-terminal triphosphate initiating the primer strand, suggesting that the interaction via the ZnFD could serve to facilitate translocation, and/or to avoid enzyme dissociation. In fact, we have observed that PrimPol primer synthesis is highly processive at least until 10–12 nt, in contrast to the distributive pattern of elongation described when starting polymerization from a pre-existing primer [21,34]. Interestingly, sustained interaction with the 5'-triphosphate of the initiating nucleotide guarantees the processive synthesis of mature primers that can be easily delivered to the replicase, due to the much higher affinity of the latter for the primer-terminus.



4. Primase and polymerase configurations of human PrimPol

The crystal structure of the AEP core of human PrimPol has been solved in complex with a template/primer and dATP as incoming 3'-site nucleotide, albeit lacking the important regions, such as the C-terminal domain containing the Zn-finger and the RBD [38] (Fig. 6). The structure revealed that the AEP core of PrimPol primarily interacted with the template DNA strand, but showed no contacts with the primer DNA strand, explaining the large distributivity displayed by PrimPol during elongation of a pre-existing primer that does not contain a 5'-terminal triphosphate [31]. We showed that the elimination of the C-terminal extension of the AEP core of human PrimPol, which removes the Zn-finger and RPA-interacting domains, stimulates of the DNA polymerase activity measured on a template/primer [23]. These results suggest that the ZnFD can interfere with the binding of a pre-existing primer, and must be mobilized (or eliminated) for efficient polymerization (Fig. 6). Thus, the current PrimPol structure mainly represents the polymerization configuration of human PrimPol (Fig. 6), and is not informative about the molecular details of the primase reaction in PrimPols, and how the C-terminal Zn-finger contributes to this activity.

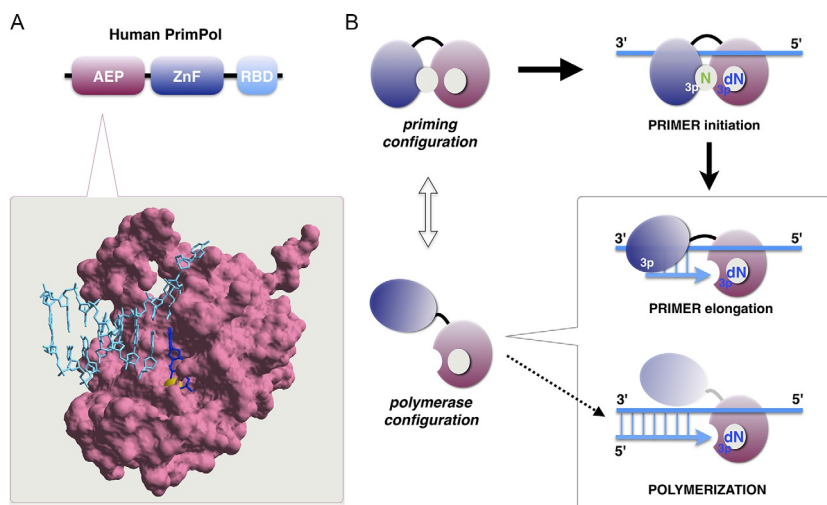


Fig. 6 Putative primase and polymerase configurations of human PrimPol, based on the mobilization of the ZnFD. (A) The image represents the crystal structure of human PrimPol AEP core forming a ternary complex with template-primer DNA (cyan) and incoming dATP (blue). The image has been generated from PDB id: [5L2X](#) (aa 1–348; [\[38\]](#)). (B) Hypothetic configuration of the ZnFD (blue) relative to the AEP core (pink) during primer initiation, primer elongation or polymerization from a pre-existing primer.

In a recent paper [\[40\]](#), Zhao and coworkers used enzyme kinetic analyses and computer simulations to dissect the mechanism by which PrimPol transfers a nucleotide to a primer-template DNA, concluding that divalent cations alter the rate-limiting step of PrimPol-catalyzed DNA elongation. According to the authors, in the presence of Mn^{2+} , a conformational transition step from non-productive to productive PrimPol: DNA complexes limits the enzymatic turnover, whereas, in the presence of Mg^{2+} , the chemical step becomes rate limiting. Both kinetic data and simulations support the notion that Mn^{2+} can be a preferred metal cofactor for PrimPol, and that PrimPol could adopt different conformations. As stressed by the authors, their conclusions can be extrapolated to the elongation stage of PrimPol-catalyzed priming, but further studies will be required to address the kinetic steps during dinucleotide synthesis and their dependence on Mn^{2+} ions.

As the ZnFD is crucial for the primase activity, specifically for interacting with the 5'-initiating nucleotide, it should adopt a “close” configuration to initiate priming. The intrinsic flexibility of the tether connecting the AEP core and the ZnFD would allow conformational changes during progressive primer elongation while maintaining the interaction with the

5'-triphosphate moiety (Fig. 6). Additional crystal structures of PrimPol in primase mode will be instrumental to evaluate how the ZnFD contributes to stabilizing the incoming 5' nucleotide, which specific residues are involved in the interaction with its sugar and phosphate residues, and whether these interactions are modulated during processive elongation to build a mature primer.



5. TLS abilities of human PrimPol are needed during DNA priming

In 2013, we described human PrimPol as the only eukaryotic primase that can initiate *de novo* synthesis of DNA. Also, we showed that PrimPol can behave as a TLS polymerase *in vitro* by elongating the pre-existing primers across DNA lesions, such as 8oxoG, or by realigning the primer ahead of replication-blocking lesions, such as abasic sites or 6-4 photoproducts [21,23]. Importantly, whereas a number of *in vivo* studies have shown that PrimPol's ability to make new DNA primers is the most important feature of PrimPol to rescue stalled forks and to maintain an optimal fork rate during both nuclear and mitochondrial DNA replication, there is no current data supporting the *in vivo* relevance of its TLS capability.

It has been published that PrimPol can read bulky UV lesions [41], but this *in vitro* data has not been supported by the structural data. The crystal structure of PrimPol in ternary complex (see Fig. 6) clearly revealed that, unlike a specialized TLS polymerase as Pol η which is endowed with the capacity to accommodate UV lesions, human PrimPol has no room for that in its active site [38]. Moreover, PrimPol active site neatly separates the templating base from the next to be copied base (by rotating the intervening phosphodiester bond), something that cannot occur with thymidine dimers as 6-4PP or CPD lesions. Conversely, human PrimPol can mobilize the primer-terminus (both back and forth), based on neighbor microhomologies, to skip bulky and unreadable lesions as 6-4PP [23,29]. Such a characteristic is not unique for human PrimPol, but a feature shared by other PrimPols, such as BcMCM, a bacterial PrimPol studied by us prior to human PrimPol [16], or TthPrimPol, a thermostable version that is being used for DNA amplification [17].

We have recently shown that PrimPol is able to bypass lesions as 8oxoG or abasic sites found immediately after initiating primer synthesis, by means of the various TLS mechanisms previously described. Strikingly, a single arginine residue, Arg²⁸⁸, although irrelevant to build *de novo* primers

across undamaged DNA, was crucial to realign the primer in order to skip unreadable lesions as abasic sites (our unpublished data). These data support the proposal that PrimPol functions as a TLS primase, as it is not only capable of making DNA primers beyond lesions, but also empowered with TLS ability to synthesize primers across template lesions.



6. AEP superfamily: Domain associations define functional classes

As indicated by Koonin, Aravind and coworkers in their seminal study [11], the AEP superfamily includes a large variety of enzymes with an RRM (RNA Recognition Motif)-like fold [9,10], that is also shared by viral RNA-dependent RNA polymerases, reverse transcriptases, cyclases and nucleic acid polymerases from the A/B/Y families, and that it likely represents a minimal nucleotidyl transferase domain. The AEP superfamily includes RNA primases, RNA/DNA primases, DNA primases (PrimPols), and some RNA polymerases involved in repair (bacterial NHEJ). The reason for this functional variability relies on the intrinsic capacity of the AEP core domain to bind incoming ribonucleotides or deoxyribonucleotides, but also on its insufficiency to stabilize a second nucleotide (primases) or a primer (polymerases) at the “5′ initiation site”. Thus, having an AEP domain does not necessarily imply being a polymerase or a primase.

As depicted in Fig. 7, conventional AEP primases are dimeric, as the AEP-containing subunit (catalytic) needs an unrelated subunit (non-catalytic) to bind the initiating 5′NTP [18]. Similarly, monomeric primases, such as PrimPols, use a specific subdomain linked to the AEP core for this function. We have shown that the Zn-finger located at the C-terminal extension of human PrimPol is functionally equivalent to the C-terminal domain of the Pri2/PriL subunit of the human Pol alpha-primase complex, being crucial to stabilize the binding of the 5′ nucleotide and enabling PrimPol to act as a primase [35]. Similarly, in bacterial PrimPols as *Tth*PrimPol, an unrelated C-terminal extension of the AEP core, named PriCT-1 domain [11,17], plays a similar role (our unpublished data).

Prokaryotic LigD is a multidomain enzyme having three enzymatic modules: a polymerase (AEP core), fused to a nuclease and a ligase (Fig. 7). In *Mycobacterium tuberculosis* LigD, this AEP core alone, named *Mt*PolDom across the literature [42,43,44,45] is able to form a so-called pre-ternary complex (PolDom:DNA:NTP) at a 3′ protruding DNA end [44] supporting the competence of the AEP-core for the binding of the incoming 3′ nucleotide. A similar pre-ternary complex, but in this case with an ssDNA template and a dNTP,

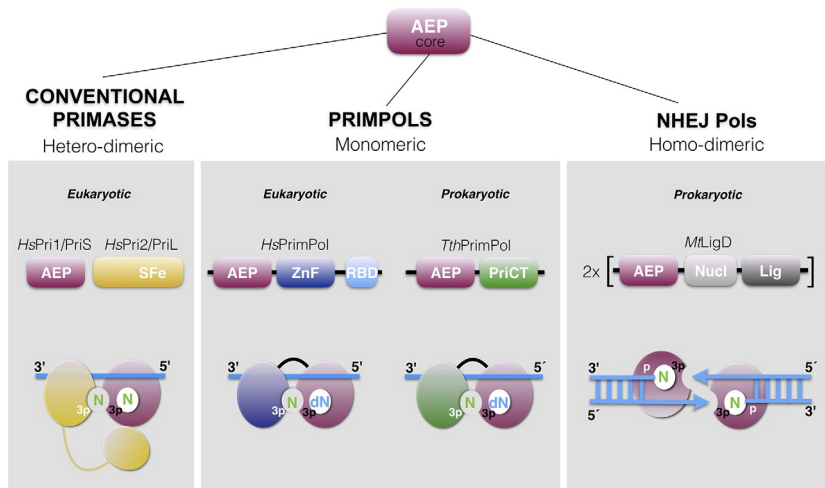


Fig. 7 Diverse associations define functional classes of AEP enzymes. Conventional AEP primases are heterodimeric, and the non-catalytic subunit (Pri2/PriL) binds the 5'NTP "primer" to form the initial dimer with a second NTP. PrimPols are monomeric, and use an extended C-terminal domain (ZnF in human PrimPol and PriCT-1 in bacterial *TthPrimPol*) to bind the 5'-NTP "primer" to form the initial dimer with a dNTP. NHEJ pols as the PolDom (AEP core) of *MtLigD* are homodimeric, and they cannot prime because they cannot use nucleotides as primers; conversely, each monomer provides the 3' protrusion of a DNA end as primer to the other monomer, promoting synapsis and NHEJ of DSBs.

was shown to be formed also with human PrimPol, even in the absence of the Zn-finger domain [35], again supporting the competence of the AEP-core for the stable binding of an incoming 3' nucleotide. In stark contrast to primases and PrimPols, *MtPolDom* cannot start "de novo" priming [44], due to its inability to bind a 5'NTP. Therefore, the AEP core of *MtLigD* (PolDom) needs to operate as a homodimer to be able to transfer the incoming 3' nucleotides, which could be required prior ligation and sealing of the DSB [43,45].

Therefore, not all AEP domain-containing proteins are primases or PrimPols, and therefore AEP enzymes should not be collectively included in a general grouping with other polymerases, in disagreement with a recent proposal [31].

References

- [1] A. Kornberg, T. Baker, DNA Replication, second ed., Freeman, San Francisco, 1992.
- [2] H. Lodish, A. Berk, S.L. Zipursky, P. Matsudaira, D. Baltimore, J.E. Darnell, Molecular Cell Biology, fourth ed., W.H. Freeman & Co, New York, 1999.
- [3] L. Blanco, M. Salas, Characterization and purification of a phage phi 29-encoded DNA polymerase required for the initiation of replication, Proc. Natl. Acad. Sci. U. S. A. 81 (1984) 5325–5329.

- [4] S. Wanrooij, J.M. Fuste, G. Farge, Y. Shi, C.M. Gustaffson, M. Falkenberg, Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro, *Proc. Natl. Acad. Sci. U. S. A.* 105 (2008) 11122–11127.
- [5] D.N. Frick, C.C. Richardson, DNA primases, *Annu. Rev. Biochem.* 70 (2001) 39–80.
- [6] R.D. Kuchta, G. Stengel, Mechanism and evolution of DNA primases, *Biochim. Biophys. Acta* 1804 (2010) 1180–1189.
- [7] M.A. Garbacz, S.A. Lujan, A.B. Burkholder, P.B. Cox, Q. Wu, Z.X. Zhou, J.E. Haber, T.A. Kunkel, Evidence that DNA polymerase delta contributes to initiating leading strand DNA replication in *Saccharomyces cerevisiae*, *Nat. Commun.* 9 (2018) 858.
- [8] L. Aravind, D.D. Leipe, E.V. Koonin, Toprim, a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins, *Nucleic Acids Res.* 26 (1998) 4205–4213.
- [9] M.A. Augustin, R. Huber, J.T. Kaiser, Crystal structure of a DNA-dependent RNA polymerase (DNA primase), *Nat. Struct. Biol.* 8 (2001) 57–61.
- [10] L. Aravind, R. Mazumder, S. Vasudevan, E.V. Koonin, Trends in protein evolution inferred from sequence and structure analysis, *Curr. Opin. Struct. Biol.* 12 (2002) 392–399.
- [11] L.M. Iyer, E.V. Koonin, D.D. Leipe, L. Aravind, Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members, *Nucleic Acids Res.* 33 (2005) 3875–3896.
- [12] J.T. Evans, D.J. Leisy, G.F. Rohrmann, Characterization of the interaction between the baculovirus replication factors LEF-1 and LEF-2, *J. Virol.* 71 (1997) 3114–3119.
- [13] R.S. Pitcher, N.C. Brissett, A.J. Doherty, Nonhomologous end-joining in bacteria: a microbial perspective, *Annu. Rev. Microbiol.* 61 (2007) 259–282.
- [14] G. Lipps, The replication protein of the *Sulfolobus islandicus* plasmid pRN1, *Biochem. Soc. Trans.* 32 (2004) 240–244.
- [15] K. Beck, G. Lipps, Properties of an unusual DNA primase from an archaeal plasmid, *Nucleic Acids Res.* 35 (2007) 5635–5645.
- [16] J. Sánchez-Berrondo, P. Mesa, A. Ibarra, M.I. Martínez-Jiménez, L. Blanco, J. Méndez, J. Boskovic, G. Montoya, Molecular architecture of a multifunctional MCM complex, *Nucleic Acids Res.* 40 (2012) 1366–1380.
- [17] A.J. Picher, B. Budeus, O. Wafzig, C. Kruger, S. García-Gómez, M.I. Martínez-Jiménez, A. Díaz-Talavera, D. Weber, L. Blanco, A. Schneider, TruePrime is a novel method for whole-genome amplification from single cells based on *Tth*PrimPol, *Nat. Commun.* 7 (2016) 13296.
- [18] A.G. Baranovskiy, N.D. Babayeva, Y. Zhang, J. Gu, Y. Suwa, Y.I. Pavlov, T.H. Tahirou, Mechanism of concerted RNA-DNA primer synthesis by the human primosome, *J. Biol. Chem.* 291 (2016) 10006–10020.
- [19] B. Arezi, B.W. Kirk, W.C. Copeland, R.D. Kuchta, Interactions of DNA with human DNA primase monitored with photoactivatable cross-linking agents: implications for the role of the p58 subunit, *Biochemistry* 38 (1999) 12899–12907.
- [20] V.B. Agarkar, N.D. Babayeva, Y.I. Pavlov, T.H. Tahirou, Crystal structure of the C-terminal domain of human DNA primase large subunit: implications for the mechanism of the primase-polymerase alpha switch, *Cell Cycle* 10 (2011) 926–931.
- [21] S. García-Gómez, A. Reyes, M.I. Martínez-Jiménez, E.S. Chocrón, S. Mourón, G. Terrados, C. Powell, E. Salido, J. Méndez, I.J. Holt, L. Blanco, PrimPol, an archaea primase/polymerase operating in human cells, *Mol. Cell* 52 (2013) 541–553.
- [22] G. Lipps, S. Rother, C. Hart, G. Krauss, A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity, *EMBO J.* 22 (2003) 2516–2525.
- [23] S. Mourón, S. Rodríguez-Acebes, M.I. Martínez-Jiménez, S. García-Gómez, E.S. Chocrón, L. Blanco, J. Mendez, Repriming of DNA synthesis at stalled replication forks by human PrimPol, *Nat. Struct. Mol. Biol.* 20 (2013) 1383–1389.

- [24] L. Wan, J. Lou, Y. Xia, B. Su, T. Liu, J. Cui, Y. Sun, H. Lou, J. Huang, hPrimpol1/CCDC111 is a human DNA primase-polymerase required for the maintenance of genome integrity, *EMBO Rep.* 14 (2004) 1104–1112.
- [25] R. Torregrosa-Muñumer, J.M.E. Forslund, S. Goffart, A. Pfeiffer, G. Stojkovic, G. Carvalho, N. Al-Furoukh, L. Blanco, S. Wanrooij, J.L. Pohjoismaki, PrimPol is required for replication reinitiation after mtDNA damage, *Proc. Natl. Acad. Sci. U. S. A.* 114 (2017) 11398–11403.
- [26] K. Kobayashi, T.A. Guillian, M. Tsuda, J. Yamamoto, L. Bailey, I. Iwai, S. Takeda, A. Doherty, J.K. Hirota, Repriming by PrimPol is critical for DNA replication restart downstream of lesions and chain-terminating nucleosides, *Cell Cycle* 15 (2016) 1997–2008.
- [27] D. Schiavone, S.K. Jozwiakowski, M. Romanello, L.J. Bailey, J.E. Sale, A.D. Doherty, PrimPol is required for replicative tolerance of G quadruplexes in vertebrate cells, *Mol. Cell* 61 (2016) 161–169.
- [28] S. Svikovic, A. Crisp, S.M. Tan-Wong, T.A. Guillian, A.J. Doherty, N.J. Proudfoot, et al., R-loop formation during S phase is restricted by PrimPol-mediated repriming, *EMBO J.* 3 (2019) 38.
- [29] M.I. Martínez-Jiménez, S. García-Gómez, K. Bebenek, G. Sastre-Moreno, P.A. Calvo, A. Díaz-Talavera, T.A. Kunkel, L. Blanco, Alternative solutions and new scenarios for translesion DNA synthesis by human PrimPol, *DNA Repair (Amst)* 29 (2015) 127–138.
- [30] M.K. Zafar, A. Ketkar, M.F. Lodeiro, C.E. Cameron, R.L. Eoff, Kinetic analysis of human PrimPol DNA polymerase activity reveals a generally error-prone enzyme capable of accurately bypassing 7,8-dihydro-8-oxo-2'-deoxyguanosine, *Biochemistry* 53 (2014) 6584–6594.
- [31] T.A. Guillian, S.K. Jozwiakowski, A. Ehlinger, R.P. Barnes, S.G. Rudd, L.J. Bailey, J.M. Skehel, K.A. Eckert, W.J. Chazin, A.J. Doherty, Human PrimPol is a highly error-prone polymerase regulated by single-stranded DNA binding proteins, *Nucleic Acids Res.* 43 (2015) 1056–1068.
- [32] M.I. Martínez-Jiménez, A. Lahera, L. Blanco, Human PrimPol activity is enhanced by RPA, *Sci. Rep.* 7 (2017) 783.
- [33] D. Branzei, B. Szakal, DNA damage tolerance by recombination: molecular pathways and DNA structures, *DNA Repair (Amst)* 44 (2016) 68–75.
- [34] B.A. Keen, S.K. Jozwiakowski, L.J. Bailey, J. Bianchi, A.J. Doherty, Molecular dissection of the domain architecture and catalytic activities of human PrimPol, *Nucleic Acids Res.* 42 (2014) 5830–5845.
- [35] M.I. Martínez-Jiménez, P.A. Calvo, S. García-Gómez, S. Guerra-González, L. Blanco, The Zn-finger domain of human PrimPol is required to stabilize the initiating nucleotide during DNA priming, *Nucleic Acids Res.* 46 (2018) 4138–4151.
- [36] A. Díaz-Talavera, P.A. Calvo, D. González-Acosta, M. Díaz, G. Sastre-Moreno, L. Blanco-Franco, G. Guerra, M.I. Martínez-Jiménez, J. Méndez, L. Blanco, A cancer-associated point mutation disables the steric gate of human PrimPol, *Sci. Rep.* 9 (2019) 1121.
- [37] P.A. Calvo, G. Sastre-Moreno, C. Perpiñá, S. Guerra, M.I. Martínez-Jiménez, L. Blanco, The invariant glutamate of human PrimPol DxEx motif is critical for its Mn²⁺-dependent distinctive activities, *DNA Repair (Amst)* 77 (2019) 65–75.
- [38] O. Rechkoblit, Y.K. Gupta, R. Malik, R.D. Rajashankar, R.E. Johnson, L. Prakash, S. Prakash, S.A.K. Aggarwal, Structure and mechanism of human PrimPol, a DNA polymerase with primase activity, *Sci. Adv.* 2 (2016) e1601317.
- [39] B. Zhu, L. Wang, H. Mitsunobu, X. Lu, A.J. Hernández, Y. Yoshida-Takashima, T. Nunoura, S. Tabor, C.C. Richardson, Deep-sea vent phage DNA polymerase specifically initiates DNA synthesis in the absence of primers, *Proc. Natl. Acad. Sci. U. S. A.* 114 (2017) E2310–E2318.

- [40] W. Xu, W. Zhao, N. Morehouse, M.O. Tree, L. Zhao, Divalent cations alter the rate-limiting step of PrimPol-catalyzed DNA elongation, *J. Mol. Biol.* 431 (2019) 673–686.
- [41] J. Bianchi, S.G. Rudd, S.K. Jozwiakowski, L.J. Bailey, V. Soura, E. Taylor, I. Stevanovic, A.J. Green, T.H. Stracker, H.D. Lindsay, A.J. Doherty, PrimPol bypasses UV photoproducts during eukaryotic chromosomal DNA replication, *Mol. Cell* 52 (2013) 566–573.
- [42] R.S. Pitcher, N.C. Brissett, A.J. Picher, P. Andrade, R. Juarez, D. Thompson, G.C. Fox, L. Blanco, A.J. Doherty, Structure and function of a mycobacterial NHEJ DNA repair polymerase, *J. Mol. Biol.* 366 (2007) 391–405.
- [43] N.C. Brissett, R.S. Pitcher, R. Juarez, A.J. Picher, A.J. Green, T.R. Dafforn, G.C. Fox, L. Blanco, A.J. Doherty, Structure of a NHEJ polymerase-mediated DNA synaptic complex, *Science* 318 (2007) 456–459.
- [44] N.C. Brissett, M.J. Martin, R.S. Pitcher, J. Bianchi, R. Juarez, A.J. Green, G.C. Fox, L. Blanco, A.J. Doherty, Structure of a preternary complex involving a prokaryotic NHEJ DNA polymerase, *Mol. Cell* 41 (2011) 221–231.
- [45] N.C. Brissett, M.J. Martin, E.J. Bartlett, J. Bianchi, L. Blanco, A.J. Doherty, Molecular basis for DNA double-strand break annealing and primer extension by an NHEJ DNA polymerase, *Cell Rep.* 5 (2013) 1108–1120.